Summary

About this Release

About this Release

The Australian Bureau of Statistics (ABS) conducts a variety of household surveys to collect data about households or the persons in them. These surveys typically use a multistage sample design that first selects a set of geographic areas and then a sample of dwellings to be approached by interviewers. Such a design gives little control over the types of households and persons that are selected - so it is important to use estimation techniques that make some correction for any imbalance in the sample. This combination of a multistage design with a possibly complex estimation technique makes variance estimation a non-trivial problem.

Within the ABS, household survey estimation has been undergoing a period of transition. In the 1980s a technique known as post-stratified ratio estimation was typically used. For these estimates a valid variance estimator was available, known in the ABS as split-halves. Standard practice was to use split-halves estimates of standard error as the basis of the standard error models published with ABS data.

Over time the drive to do more with each survey has increased the complexity of sample design and required extensions to estimation. Now weighting may include a separate non-response adjustment phase, and may use a variety of auxiliary data - for example, population counts for households as well as persons. The enabling technology for some of these extensions was the CALMAR macro of Deville, Särndal and Sautory (1993), which allows weighting using the generalised regression method and related calibration methods..

In the ABS these changes to estimation methods moved faster than the ability to estimate variances. Up until 1997 split-halves was the only tool used for variance estimation, and any bias in the resulting standard error models was not measured. In late 1997 this lack was identified as a key problem, and it was intensively researched over a two year period. As a result the ABS is moving to a new standard for weighting and for variance estimation.

This paper describes the methodological principles behind the generalised regression weighting approach used in the ABS for estimation in household surveys. It then presents a number of approaches to estimating variance for such estimates, and recommends the group jackknife approach for standard use in the ABS. These methods are available via SAS macros written within the ABS. They will be implemented as components of the Household Survey Facilities (HSF) processing system. It is also planned to make the group jackknife approach available through the SUPERCROSS tabulation and aggregation facility so that variance estimates are available for ad hoc requests.

Section 2 describes the development of weighting techniques and in particular the generalised regression weighting approach. Section 3 presents a variety of methods for estimating the variance of these estimates. Section 4 focuses specifically on the group jackknife approach to variance estimation, and presents a theoretical justification. Section 5 compares the variance estimators in the context of a simulated population with systematic sampling of clusters. Section 6 describes an actual survey, the Australian Labour Force Survey, and presents a comparison of different variance estimators. It also shows how the new methods apply to estimating variance for complex estimates such as the trend. Section 7 summarises results from a number of evaluations conducted on more complex surveys. These evaluations focus on the value of the estimates for deriving variance models. Section 8 concludes with some comments on the application of the techniques in generalised facilities in the ABS